DOCUMENT RESUME

ED 234 364 CS 007 300

AUTHOR Rupley, William H.; Wise, Beth S.
TITLE Reading Teacher Effectiveness Research: Generalizability of Significant Findings. Instructional Research Laboratory Technical Series #R83002.
INSTITUTION Texas A and M Univ., College Station. Instructional Research Lab.
PUB DATE Sep 83
NOTE 16p.
PUB TYPE Information Analyses (070)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Classroom Research; Data Analysis; Educational Research; Elementary Education; Reading Research; *Reading Teachers; Reliability; *Research Design; *Research Methodology; *Research Problems; Science History; *Teacher Effectiveness; Validity
IDENTIFIERS Generalizability Theory; Research Trends

ABSTRACT

Research on reading teacher effectiveness has taken several different directions over the past 30 years: during the 1950s and most of the 1960s, research focused primarily on teacher qualities; in the early 1970s attention shifted to the effect of the teaching process on student learning, while in the late 1970s and early 1980s experiments defined more specifically factors involved in teacher effectiveness. Experimental design as well as means of data collection and analysis also altered during this time. The isolated classroom observation of the 1960s, for example, was replaced in the 1970s by more direct observation in classroom settings. Despite distinct improvements, such as refinements in the determination of dependent and independent variables in teaching effectiveness, the generalizability of significant findings continues to be limited by methodological and experimental design problems. (MM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED234364

Reading Effectiveness Research: Generalizability of Significant Findings

William H. Rupley
Texas A&M University

Beth S. Wise
McNeese State University

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

William H. Rupley

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Instructional Research Laboratory
Educational Curriculum & Instruction
Texas A&M University, College Station, TX.
Technical Series # R83002

September, 1983

CS 007300

William H. Rupley
Associate Professor
Texas A&M University
Department of EdCI
College Station, TX 77843

Reading Teacher Effectiveness Research: Generalizability of Significant Findings

William H. Rupley
Texas A&M University

Beth S. Wise
McNeese State University

Historically, research in effective teaching has taken several different directions over the past 30 years. Within this span of inquiry, three distinct time periods are identifiable: (1) 1950s and 1960s, (2) early 1970s, and (3) mid 1970s to the present (Duffy, 1980). Collectively, many of the research findings from these time periods can be linked with effective reading instruction because student reading achievement often served as the dependent variable (Centra and Potter, 1980). Although such research has expanded the knowledge base in reading teacher effectiveness, there are several issues related to variable specification and generalizability of findings that need to be considered in both the extant literature

and future research (Shavelson and Russo, 1977).

Major threats to the generalizability of research findings in reading teacher effectiveness can come from a variety of sources; however, those that appear to be pervasive throughout the past 30 years are related to variables under investigation, data gathering procedures, and analyses.

Mid 1950s and 1960s

The primary focus of teacher effectiveness research during the 1950s and most of the 1960s was on the qualities of the teacher. A quality that was extensively investigated was teachers' personality. Getzel and Jackson (1963) reported that over 1000 studies of teacher personality had been conducted in the late forties and early fifties. They concluded that after more than fifty years of inquiry effort little is known about the effects of teacher personality and effective instruction.

In addition to using personality as an independent variable, the relationship of other teacher characteristics and instructional methods to students' achievement were also explored. Sex, education, race, and years of experience are examples of independent variables that researchers' attempted to link to effective teaching.

Researchers used a variety of procedures to gather data on teacher characteristics, instructional methods, student characteristics and student achievement. Questionnaires, surveys, rating scales, and observation instruments were administered to teachers, supervisors, and students in a search for significant relationships with student achievement. Classroom observations were used to study

predetermined teacher behaviors assumed to be associated with student achievement. Usually these observations focused on verbal interactions using Flanders Interaction Analysis Categories (Flanders, 1960) and classroom variables dealing with emotional climate, verbal emphasis, and student-initiated activity employing instruments such as the Observation Schedule and Record (Medley and Mitzel, 1958).

An investigation of elementary reading during the mid sixties also used direct classroom observation to study teachers' implementation of a single method of teaching beginning reading (Chall and Feldman, 1966). Students' reading gains served as the dependent variable. Data gathered from the direct observations and questionnaires were the independent variables.

Findings from the majority of these research investigations lacked external validity. Critics of this research attacked it as being isolated and remote from the actual classroom (Cogan, 1963). Wallen and Travers (1963) felt that for progress to go forward, theory should precede practice in teacher effectiveness research; however, their call went unanswered well into the late 1960s.

Although in the late sixties some attention was given to gathering data in naturalistic classroom settings, data were often limited to behaviors that were part of the content of the instrument used, which did not allow for the systematic recording of classroom events that occurred outside of the specified content. Furthermore, minimal attention, if any, was given to reliability and validity of observation systems (Rupley and Mangano, 1982) and other data gathering procedures, such as questionnaires and rating scales. Teachers' reports of classroom instruction and supervisors' ratings of teachers'

effectiveness were assumed to be accurate indicators of what actually occurred in classroom reading instruction. Such methodological flaws resulted in a major threat to the generalizability of any significant findings and offered little application of results to either preservice or inservice teacher training programs.

Early 1970s

Several important developments occurred during the early seventies that provided a more cohesive direction for the study of teacher effectiveness. Major reviews of past research were conducted (Dunkin and Biddle, 1974; McNeil and Popham, 1973; Rosenshine and Furst, 1973) which changed the direction of research focus. One major change was a focus on the process of teaching in relation to its effect on the product, which was students' learning.

Another significant event that reshaped the focus of teacher effectiveness research was the funding of a number of major investigations between 1972 and 1975 by The National Institute of Education. Among the funded investigations were those that focused on (1) effective education of disadvantaged children (Soar and Soar, 1972), (2) stability of teacher effectiveness (Brophy and Everston, 1974), and (3) specification of effective teaching behaviors (Berliner, 1975).

Methodological features of teacher effectiveness research were also reconceptualized. Pretesting and posttesting to determine students' adjusted mean achievement in the basic skill areas of math and reading were being used as the dependent variables. Independent variables related to students' adjusted mean achievement

were teachers' instructional behaviors, students' behaviors and students' socio-economic standing. Data collection was becoming more classroom oriented, and more direct observation in natural settings was being used to record instructional activities, students' behavior, and classroom environments.

Taken as a group, the studies conducted in the early 1970s varied in the types of teachers and students included, the kinds of variables addressed, and the methods used. There was, however, replication of findings in some of the studies, even though many of these were more poorly designed than others. Correlation analyses was the primary means of data analysis and the findings were in the middle ranges (Duffy, 1980).

Generalizability of significant results, however, was considerably limited due to major methodological flaws. Data gathering procedures, although more classroom focused than in the preceding time period, still lacked major consideration being given to reliability and validity issues. Observer agreement was the only reliability issue addressed and observations in some studies were limited to only one or two episodes (Rupley and Mangano, 1982). Consistency across studies did help lend validity to the findings, however, Rosenshine (1977, 1978) cautioned against implementing these initial findings into teacher training programs before validity had been established. Finally, a serious threat to external validity was the statistical analysis employed in the majority of inquiries. With a range of 100 to 1000 measures in a single study, many significant correlations were obtained by chance. Furthermore, some

investigators applied significantly more measures in their analyses than were actually studied, which violated the assumptions underlying the statistical tests they used (Centra and Potter, 1980).

Late 1970s to present

Several major reviews of the process-product research conducted in the late sixties and early seventies helped to further refine and define the direction of teacher effectiveness research (Rosenshine, 1977, 1978; Medley, 1977; Brophy, 1979). A notable outcome was the more precise identification of the factors under investigation. Specification of independent variables such as teacher-directed instruction, pupil engagement, classroom management, and so forth became more common across investigations.

Another major thrust was the emergence of classroom based experimental studies of teacher effectiveness. These experimental efforts were designed to test the validity of results of the large scale correlation studies conducted in the early seventies (Anderson, Everston, and Brophy, 1979; Good and Grouws, 1977; Stallings, Needels and Stayrock, 1979). This new experimental focus had two major stages (Gage and Giaconia, 1981). The first stage consisted of training one group of teachers to employ process variables associated with effective instruction and withholding training from a comparable group of teachers. Teachers' use of process variables were measured through direct observation in the teachers' classroom. The second stage was characterized by using observation data as the independent variables and students' product outcome, such as reading achievement, as the dependent variable.

The experimental focus of teacher effectiveness research holds considerable promise for more accurate specification of instructional processes that cause student learning. However, the findings from the major studies using such a methodology still have limited generalizability and are open to question about causal influences (Anderson, 1979).

Factors under observation were not uniformly defined, thus, what was student engagement in one study may have been coded as a different behavior in other studies. Reliability was most often addressed in terms of inter-observer agreement, which only provides a coefficient for agreement between observers. For example, a major violation occurring frequently was to train observers until they reached an established criterion for major categories which fails to account for the range of variation of each behavior within that category.

Little attention was given to establishing generalizability coefficients for each subcategory of an observation system. Generalizability theory assumes that the sample is equivalent to a set of possible combinations of the conditions for which observations can be made. Observations made within a particular facet are generalizable to other similar situations. The purpose of establishing generalizability is to determine the degree of variability for each facet. Without researchers addressing this issue in the development and use of their observation systems, significant findings will continue to have limited generalizability.

Inappropriate or weak experimental designs and data analyses continue to limit the external validity of recent teacher effectiveness research findings. Groups in some inquiries are not comparable

nor adequately described. Data were in some studies analyzed by discrete behaviors for each major category, but observers' agreement was determined by their reaching criterion for the overall major category. Finally, statistical tests also were violated. Fifty-five one way analyses of variances were conducted in one inquiry (Anderson, Everston, and Brophy, 1979) and significant process variables at the .05 level were reported. In this instance, significant p values would have had to be equal to or less than .0009.

Summary

As noted in Figure 1, major changes have occurred in the factors investigated and the data gathering procedures employed in reading teacher effectiveness research. The generalizability of significant findings continues to be limited by methodological and experimental design problems. However, it has been suggested (Good, 1979) that outcomes from experimental studies will not ever be predictable, since several teacher behaviors could be used to create the same effect, and identical teachers could have different impacts on different students. Anderson (1979) has suggested that since it is impossible to control for all the variables, the classroom researcher must somehow reach a compromise between scientific theory and classroom reality.

Admittedly, a compromise is a fact of classroom research; however, a compromise of the research tools - design, data gathering procedures, and analyses - should not be major threats to the generalizability of significant results. The importance of researchers giving careful attention to these research tools becomes even more important when the magnitude of teacher effects on student reading achievement

is considered. McDonald's research (1976) has attributed 36 percent of the variance of students' end-of year reading achievement to teacher effects; therefore, the magnitude of effect of a single process variable on student achievement is going to be extremely small.

Insert Figure 1 here

Figure 1: Major threats to the generalizability of research findings in reading teacher effectiveness research in relation to factors investigated and data gathering procedures utilized during selected time periods.

Time Periods	Factors Investigated	Data Gathering Procedures	Major Threats to Generalizability
Late 1950s & 1960s	D--Student Attitudes Student Achievement I--Teacher Personality Educational Characteristics Teacher Instruction Methods/materials	Self-Reports & Questionnaires Supervisor's Rating Scales Students' Rating Scales Teacher's Perceptions Classroom Observation of Predetermined Teacher Behaviors (usually verbal interactions and personality factors)	Methodological--lack of reliability and validity of data gathering procedures and inappropriate experimental designs. Variable specification--lack of attention to validity of independent variables.
Early 1970s	D--Student Attitudes Student Adjusted Achievement I--Teacher Instruction Student Behavior SES	Self-reports & Questionnaires Classroom Observation of Behavior	Methodological--lack of reliability and validity of data gathering procedures and inappropriate or weak experimental designs. Variable Specification--lack of attention to validity
Mid 1970s to Present	D--Student Attitudes Student Adjusted Achievement I--Teacher Instruction Student Behavior SES School Classroom	Classroom Observation of Predetermined Teacher and Student Behavior in Instruction Classroom Observation That is Descriptive of Events and Observer's Impressions	1) Methodological--lack of careful attention to reliability and validity of data gathering procedures and weaknesses in experimental designs and analyses.

References

Anderson, L. M. Classroom-based experimental studies of teaching effectiveness in elementary schools. R & D Report No. 4102, Austin, TX: Research and Development Center for Teacher Education, University of Texas, 1979.

Anderson, L., Everston, C., & Brophy, J. An experimental study of effective teaching in first grade reading groups. Elementary School Journal, 1979, 79, 193-223.

Berliner, D.C. The beginning teacher evaluation study: Overview and selected findings, 1974-1975. A paper presented at the National Invitational Conference on Research on Teacher Effects: An Examination by Decision-Makers and Researchers, Austin, Texas, 1975.

Brophy, J., & Everston, C. Process-product correlations in the Texas effectiveness study: Final report. R & D Report No. 74-4, Austin, TX: Research and Development Center for Teacher Education, University of Texas, 1974.

Brophy, J.E. Advances in teacher effectiveness research. Occasional paper no. 18. National Institute of Education, Washington, D.C., April, 1979.

Centra, J.A., & Potter, D.A. School and teacher effects: An interrelational model. Review of Educational Research, 1980, 50, 273-291.

Chall, J., & Feldman, S. First grade reading: An analysis of the interactions of professed methods, teacher implementation and child background. The Reading Teacher, 1966, 19, 569-575.

Cogan, M.L. Research on the behaviors of teachers: A new phase. Journal of Teacher Education, 1963, 14, 238-243.

Duffy, G. Teacher effectiveness research: Implications for the reading profession. A paper presented at the annual meeting of the National Reading Conference, San Diego, 1980.

Durkin, M.J., & Biddle, B.J. The study of teaching. New York, NY: Holt, Rinehart, and Winston, 1974.

Flanders, N.A. Teacher influence, pupil attitudes, and achievement. Minneapolis, MN: University of Minnesota, 1960.

Gage, N.L., & Giaconia, R. Teaching practices and student achievement: Causal connections. New York University Quarterly, 1981, Spring, 2-9.

Getzel, J.W., & Jackson, P.W. The teacher's personality and characteristics. In N.L. Gage (Ed.): Handbook of research on teaching, Chicago, IL: Rand McNally & Company, 1963, 506-582.

Good, T.L. Teacher effectiveness in the elementary school. Journal of Teacher Education, 1979, 30, 52-64.

Good, T.L., & Grouws, D.A. Teaching effects: Process-product study fourth grade mathematics classrooms. Journal of Teacher Education, 1977, 28, 49-54.

McNeil, J.D., & Popham, W.J. The assessment of teacher competence. In R.M.W. Travers (Ed.): Second handbook of research on teaching, Chicago, IL: Rand NcMally Publishing Company, 1973.

Medley, D. Teacher competency and teacher effectivness: A review of process-product research. Washington, D.C.: American Association of Colleges for Teacher Evaluation, 1977.

Medley, D.M., & Mitzel, H.E. A technique for measuring classroom behavior. Journal of Educational Psychology, 1958, 49, 86-93.

Rosenshine, B.V. Academic engaged time, content covered, and direct instruction. A paper presented at the annual meeting of the American Educational Research Association, New York, 1977.

Rosenshine, B.V. Instructional principles in direct instruction. A paper presented at the annual meeting of the American Educational Research Association, Toronto, 1978.

Rosenshine, B.V., & Furst, N. The use of direct observation to study teaching. In R.M.W. Travers(Ed.): Second handbook of research on teaching, Chicago, IL: Rand McNally Publishing Company, 1973, 122-183.

Rupley, W.H., & Mangano, N.G. Development and measurement issues associated with reading classroom observation systems. In J.A. Niles & L.A. Harris (Eds.): New inquiries in reading research and instruction, Rochester, NY: The National Reading Conference, 1982, 200-203.

Shavelson, R., & Russo, N.A. Generalizability of measures of teacher effectiveness. Educational Researcher, 1977, 19, 171-183.

Soars, R.S., & Soar, R.M. An empirical analysis of selected Follow Through Programs: An example of a process approach to evaluation. In I. Gordon (Ed.): Early childhood education, Chicago, IL: National Society for the Study of Education, 1972, 229-261.

Stallings, J., Needels, M., & Stayrock, N. The teaching of basic reading skills in secondary schools. Phase I and Phase II. SRI International, Menlo Park, CA, 1979.

Wallen, N.E., & Travers, R.M.W. Analysis and investigation of teaching methods. In N.L. Gage (Ed.): Handbook of research on teaching, Chicago, IL: Rand McNally & Company, 1963, 448-505.